Harvard Data Science Review

The Mean-ness of Statistics

Arunabh Ghosh[1]

[1]Department of History, Harvard University, Cambridge, Massachusetts, United States of America

Published on: Oct 27, 2022

DOI: https://doi.org/10.1162/99608f92.3e79086d

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

Column Editors' Note: Statistical theory, like numbers themselves, superficially might seem to travel seamlessly across linguistic, political, and economic divides. But as Ghosh's column demonstrates, what counts as statistically 'valid' is deeply tied to social and political circumstances. Using a case taken from his recent book, *Making It Count: Statistics and Statecraft in the Early People's Republic of China*, Ghosh shows us that even something as seemingly simple as enumerating and averaging is always embedded in competing political ideas about how and what to count.

Keywords: statistics, data, central tendency, China, index numbers, probability

---

Irving Fisher (1867–1947) could scarcely have imagined that a few short years after his death he would occupy centerstage in a series of heated debates in China. A brilliant econometrician, Fisher is a seminal figure in the mathematization of the discipline of economics. In December 1930, he had been among the cofounders of the Econometrics Society and served as its first president until 1934.[1] During his lifetime, the closest he came to China was through his doctoral advisee Franklin L. Ho (He Lian, 1895–1975). An influential economist in his own right, Ho had returned to China after graduating from Yale in 1926 and helped set up at Nankai University in Tianjin, one of the first research institutes dedicated to the study of China's economy. In his memoirs, he would recall fondly his time at Yale and the support he had received from Fisher. [2]

The Chinese statisticians and economists who participated in the debates during the 1950s did not, however, have praise or a belated valediction on their minds. Instead, as a conference that took place in Shanghai in May 1955 made clear, their goal was to criticize precisely those things that Fisher had devoted a lifetime to: the mathematization of statistics and economics (Tang, 1956). A key feature of this mathematization within statistics (and by extension economics too) was an embrace of probability theory and probabilistic methods (Hacking, 1990; Porter, 1986). This was a process that had long been in the making but had accelerated at the end of the 19th century under the influence of figures like Francis Galton and Karl Pearson, who devised and employed increasingly sophisticated mathematical methods to study a whole range of social phenomena. As has been well documented by historians, these methods ultimately transformed fields from psychology to sociology, and from economics to medicine.[3] Today, it is easy for us to forget that that this transformation was anything but inevitable, and in settings like that of the early People's Republic of China (PRC) and the Soviet Union, actively resisted. In what follows, I use the example of debates about the central tendency of data to explore key aspects of the Chinese critique.[4] As we shall see, the problems they grappled with are not easily dismissible as marginal or pseudoscientific oddities. Rather, they serve as a useful reminder that the acceptance of even the most basic statistical ideas and concepts is contingent and contextual. Historians have shown this to be the case for many areas of science—what counts as heresy in one decade can become textbook science in another, and vice versa. Statistics and data science, in this sense, are far from exceptional.

For the men who gathered in Shanghai in the summer of 1955, any commitment to making statistics probabilistic and economics mathematical was anathema. It was not a simple matter of ignorance or anti-elitism: many had trained extensively in both disciplines, and some also held advanced degrees from Western institutions. Instead, they rested their arguments on ontological and epistemological grounds. They argued that the natural and social world were distinct and therefore demanded distinct methods of analysis. Although chance and randomness may well be features of the natural world, they had no role to play in the study of social phenomena. They then proceeded to define statistics explicitly and exclusively as a social science, meant to study only the social world. Those who wanted to study the natural or the physical world were free to do so, but in departments of physics, mathematics, chemistry, and so on, where they could use mathematical statistics.

Such a definition of statistics as a social science was based on a specific extrapolation of Marxian theories of social change: the various stages society would traverse—primitive communism, slave society, feudalism, capitalism, socialism, and eventually, stateless communism—were known and therefore neither random nor uncertain. Therefore, no method that incorporated chance or randomness was suitable to study society. Articulated in the Soviet Union in the 1930s, this definition grew in prominence over the years and was formally ratified at a major conference in Moscow in March 1954 (Ghosh, 2020, pp. 67–71). Its rise to dominance in China can be traced to 1949, when the Communist Party of China (CCP) defeated Chiang Kai-shek's Nationalist government and established the People's Republic of China. The CCP immediately set about remodeling state and society along Marxian principles, drawing heavily upon Soviet guidance and experience in that process.[5]

In practical terms, this definitional distinction meant the rejection of probability theory and consequently of any method of data collection or analysis that relied upon probabilistic methods. Most tellingly it entailed the refusal to employ large-scale random sampling, then among the most cutting-edge technologies of data collection.[6] As a result, the principal and preferred mode of data collection became exhaustive enumeration. The census method, in other words. During the 1950s, the Chinese built an extensive periodical reporting system to carry out such data collection across all sectors of economy and society.[7] This system impressed many observers, including the Indian statistician P. C. Mahalanobis (1893–1972), who acknowledged that it was "appreciably better than that in India in respect of coverage, availability, and accuracy of data required for purposes of planning and current policy decisions" (Ghosh, 2020, p. 47). Mahalanobis was among a small number of statisticians who visited both the PRC and the Soviet Union in the 1950s. In a letter to one of his colleagues, he summarized what he saw: "one orthodox viewpoint in USSR has been to exclude mathematics from economics; to insist on statistics to be almost exclusively the handmaiden of economics; and to separate mathematical statistics as a separate (and somewhat abstract) subject under probability and mathematics" (Ghosh, 2020, p. 232; Ghosh, 2016, p. 74). The PRC, as Mahalanobis discovered during his visit in the summer of 1957, had largely followed the Soviet example. At the same time, resistance to random sampling was hardly limited to the PRC and the Soviet Union. As late as the early decades of the 20th century, many statisticians in Germany and Italy also regarded statistics largely as a social science. Although they did not employ the Marxian lens favored by the Soviets and later the Chinese, they did resist the mathematization of the discipline. In the United States, the U.S. Department of Agriculture continued to rely on exhaustive enumeration to collect agricultural data into the 1930s. And it was only with the 1940 Census that probability sampling was first used for a major statistical exercise (Bouk, 2022).

Rejecting probabilistic methods and relying upon exhaustive enumeration was not, however, the extent of the discomfort that Chinese statisticians felt with mathematical statistics. They went to great lengths to point out that they did not dispute the relevance of specific mathematical techniques, but that these techniques had to be suitable for a socialist society. Here they took special umbrage at Fisher's privileging of formal mathematics at the expense of context, understood in particular through the salience of real-world economic and political conditions. They pointed out that bourgeois statisticians such as Fisher frequently neglected to pay attention to the importance of grouping, bias, and weights when performing their calculations. In a criticism redolent with contemporary significance, they declared that calculations of per capita gross domestic product in the West did not differentiate across the population, thereby eliding disparities between the bourgeoisie (the haves) and the proletariat (the have-nots) and producing a misleading picture of economic realities on the ground (Xu & Liu, 1955, p. 33). In making these criticisms, Chinese statisticians also borrowed from Western critics of Fisher, such as the Austrian statistician Wilhelm Winkler (1884–1984).[8]

The Chinese critique of mathematical formalism typically began by demonstrating the relationship between three measures of central tendency—the arithmetic mean (AM), the geometric mean (GM), and the harmonic mean (HM).[9] As the statistician Jianzhen Wang (1955, p. 41) noted in a 1955 article, for any given set of numbers, the following relationship held true:

$$\text{AM} > \text{GM} > \text{HM}.$$

He offered a simple example to substantiate the point. Given the three numbers—400, 300, and 200—one could calculate the three mean values thus:

$$\text{Arithmetic Mean (AM)} = \frac{(400+300+200)}{3} = 300$$

$$\text{Geometric Mean (GM)} = \sqrt[3]{(400x300x200)} = 286$$

$$\text{Harmonic Mean (HM)} = \frac{3}{\left(\frac{1}{400}+\frac{1}{300}+\frac{1}{200}\right)} = 277$$

Fisher had himself summarized this relationship in a paper at the 82nd Annual Meeting of the American Statistical Association in 1920, where he noted, "as to bias, it can be shown that the arithmetic average has an upward bias, i.e., has in the very nature of the arithmetic process a natural tendency to give results too large, that the harmonic has a downward bias, that the geometric, median, and mode have no inherent bias in either direction" (Fisher, 1921, p. 536). To Chinese statisticians, the resultant preference in the West for the GM, because it putatively had no 'inherent bias,' was nothing but mathematical formalism run amok. And hiding behind this mathematical formalism, they surmised, was the desire of ‘capitalist statisticians’ such as Fisher to make price fluctuations appear less conspicuous and thereby cover up the inevitability of economic crises within capitalism (Jin, 1957, p. 30; Tang, 1956, p. 13; Wang, 1955, p. 41).

Table 1. An Economy with three types of products (Wang 1955, p. 41).

A	B	C	D
Product	**Quantity (Tons)**	**Price/ton (Yuan)**	**Share (%)**
good1	500	400	20
good2	1,000	300	30
good3	2,500	200	50

In order to communicate the critique more meaningfully, they frequently offered an example with economic content. Wang (1955, p. 41) is again useful here. In the same essay, he asked his readers to consider an economy with three types of products {good1, good2, good3} that have production and pricing characteristics listed in columns B and C (see Table 1).

Now what would a combined average price look like? As Wang explained, only one method provided the answer: the arithmetic mean:

$$\text{Average Price} = \frac{\sum BC}{\sum B}$$

$$\text{Average Price} = \frac{(500x400)\ +(1000x300)\ +(2500x200)}{(500+1000+2500)}$$

$$= 250 \text{ Yuan}$$

But what if the statistician did not have information on quantity, but only on individual prices and their relative share in monetary terms (i.e., columns C and D)? In that case, Wang explained that a combined average price could be arrived at by using the harmonic mean:

$$\text{Average Price} = \frac{\sum D}{\sum \frac{D}{C}}$$

$$\text{Average Price} = \frac{(20+30+50)}{(\frac{20}{400}+\frac{30}{300}+\frac{50}{200})}$$

$$= 250 \text{ Yuan}$$

The point Wang wanted to make was that *in no case was the geometric mean the correct choice*. This kind of analysis and the resultant dismissal of the GM became standard fare in major statistics textbooks produced during the 1950s. In those instances where the GM was taught, it came with the proviso that its use was restricted to special circumstances (Xu et al., 1956, p. 113; Zhang & Zou, 1953, pp. 41–51). As Daiguang Hu , another statistician involved in the 1955 gathering, explained, the GM could be used to calculate mean growth rates when individual rates represented change in the same direction (i.e., as part of a series that trended consistently up or consistently down) (Hu, 1955, pp. 26–27). The vast majority of statistical analysis on rates of change produced in 1950s China thus tended to favor the AM.

These preliminary discussions on measures of central tendency were typically followed by a broader critique of Western index number theory. Here, Fisher's extensive work on index numbers, in particular his life-long attempt to devise an ideal index formula (in Chinese, *lixiang gongshi*), provided grounds for a substantive attack on his legacy. Such an ideal index would be free from the bias (upward or downward) that plagued all existing indices. In 1921, Fisher proposed "two supreme tests" to help devise such an ideal index, explaining that it "should work both ways as to the two factors, prices and quantities"; and second, that "the formula should work both ways as to time" (Fisher, 1921, 1922). In mathematical terms, Fisher's ideal index formula took the form of the square root of the product of the Paasche index and the Laspeyres index.[10] In the minds of Chinese statisticians like Jianzhen Wang, Daiguang Hu, and many others, Fisher's ideal index formula ended up conflating two distinct problems, one related to the type of mean, and the other related to the types of weights. By essentially privileging mathematical principles, Fisher was deemed guilty of completely ignoring the economic content of the various indices.

The resolution of these debates in favor of 'socialist' statistics had a lasting impact on statistical practice in the PRC, albeit with brief interludes. Toward the end of the 1950s, this dominance was threatened by two contrasting experiments. Driven by the impracticality of generating reliable agricultural data using exhaustive enumeration, the first experiment involved a series of exchanges with statisticians from the Indian Statistical Institute in Calcutta, from whom Chinese statisticians hoped to learn about the latest random sampling techniques. The second detour, inspired by Mao Zedong's own theorization of social research methodology, was influential during the Great Leap Forward (1958–1962) and rejected both exhaustive and probabilistic methods in favor of typical or purposive sampling.[11] By the early 1960s, however, socialist statistics had returned to its position of theoretical and practical dominance. And despite the numerous upheavals of the ensuing years, most notably during the early phase of the Cultural Revolution (1966–1969), it would retain that status into the late 1970s.

In the end, although the Chinese critique of Fisher's ideal index was not without merit, the alternatives offered remained limited to straightforward uses of the arithmetic mean. And while this simplified the production of statistical data and made it more easily accessible, its limited scope also left generations of statisticians in China (and in the Soviet Union) poorly exposed to the richness of the discipline. They also became increasingly distanced from their colleagues in departments of mathematics, where mathematical statistics continued to thrive. The experiences of Baolu Xu (1910–1970) and Andrey Kolmogorov (1903–1987) serve as telling examples. Among the leading probabilists of their time, neither was subjected to the criticisms described here simply because they were not considered statisticians (Ghosh, 2020, p. 66n39, pp. 122–123). The siloed nature of statistical research in the PRC and in the Soviet Union stands in stark contrast to other parts of the world, where the dichotomy between techniques/concepts and disciplinary boundaries played out quite differently. In North America, Europe, and India, by contrast, there was far less friction between the statistical methods developed for mathematics, physics, and chemistry and those for the social sciences.

In the PRC, the mutual isolation between the social science of statistics and mathematical statistics only began to fade in the 1980s. The reassessment was facilitated in part by the changed political and ideological milieu of Deng Xiaoping's policy of Reform and Opening Up, which afforded new space for experimentation across the sciences. In statistics, this more ideologically relaxed environment enabled a reassessment of the core principles that had informed statistical work since the 1950s. Even so, it would take several years before the disciplinary distinction was rendered invisible.

From our contemporary vantage point, the seeming universal acceptance of statistical and data science machinery might hinder us from understanding a moment like this as something other than a mere curiosity. And yet, perhaps its most lasting legacy might be to remind us that ideology can never truly be divorced from knowledge practices. What we count and how we count are never self-evident choices. They are invariably mediated by belief-systems and disciplinary practices, with implications for both those who count and those who are counted.

Disclosure Statement

Arunabh Ghosh has no financial or non-financial disclosures to share for this article.

References

Bjerkholt, O. (2017). On the founding of the econometric society. *Journal of the History of Economic Thought*, *39*(2), 175–198. http://doi.org/10.1017/S105383721600002X

Bouk, D. (2022). *Democracy's data: The hidden stories in the U.S. Census and how to read them*. MCD; Farrar, Straus and Giroux.

Daston, L. (1988). *Classical probability in the Enlightenment.* Princeton University Press.

Fisher, I. (1921). The best form of index number. *Quarterly Publications of the American Statistical Association*, *17*(133), 533–551. https://www.tandfonline.com/doi/abs/10.1080/15225445.1921.10503812

Fisher, I. (1922). *The making of index numbers: A study of their varieties, tests, and reliability*. Houghton Mifflin.

Ghosh, A. (2016). Accepting difference, seeking common ground: Sino-Indian statistical exchanges 1951–1959. *BJHS Themes*, *1*, 61–82.

Ghosh, A. (2018). Lies, damned lies, and (bourgeois) statistics: Ascertaining social fact in midcentury China and the Soviet Union. *Osiris*, 33(1), 149–168. http://doi.org/10.1086/699237

Ghosh, A. (2020). *Making it count: Statistics and statecraft in the arly People's Republic of China*. Princeton University Press.

Gigerenzer, G., Swijtink, Z. G., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life*. Cambridge University Press.

Hacking, I. (1990). *The taming of chance.* Cambridge University Press.

Ho, F. L. (1930). *Index numbers of the quantities and prices of imports and exports and of the barter terms of trade in China, 1867–1928*. Chihli Press.

Ho, F. L. (1966). Reminiscences of Ho Lien (Franklin L. Ho), oral history, 1966, in Chinese Oral History Project, Rare Book and Manuscript Library, Columbia University, New York.

Hu, D. (1955). Ganqing zichan jieji tongji sixiang de liudu [Eliminate the pernicious influence of bourgeois statistical ideology]. *Tongji gongzuo tongxun*, *10*, 23–27.

Jin, G. (1957). Dui "Tongjixue dagang" de ziwo pipan [Self-criticism of *An outline of statistics*]. *Tongji gongzuo*, *3*, 28–32.

Kruger, L., Daston, L. J., & Heidelberger, M. (Eds.) (1987). *The probabilistic revolution* (Vol. 1) MIT Press.

Porter, T. (1986). *The rise of statistical thinking, 1820–1900*. Princeton University Press.

Porter, T. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Schmetterer, L. (1984). Tribute to Wilhelm Winkler at his 100th anniversary. *International Statistical Review / Revue Internationale de Statistique*, *52*(3), 227–228.

Tang, Q. (1956). Zichan jieji tongji zhishu lilun pipan [Criticism of the bourgeois theory of index numbers]. *Caijing yanjiu*, *1*, 13–19.

Xu, Q., & Liu, X. (1955). Guanyu zichan jieji tongji lilun de pipan [Criticisms regarding bourgeois statistical theory]. *Tongji gongzuo tongxun*, *8*, 28–34, 21.

Xu, Q., Jiang, Z., Lin, F., Zheng, Y., Liu, X., & Zhou, F. (Eds.) (1956). *Tongji lilun yiban wenti jianghua* [Common problems in statistical theory]. Tongji chubanshe.

Wang, J. (1955). Dui zichan jieji zhishu "lilun" de pipan [Criticism of the bourgeois theory of index numbers]. *Tongji gongzuo tongxun*, *11*, 40−43.

Winkler, W. (1954). Older and newer ways of solving the index numbers problem. *Bulletin of the International Statistical Institute*, *34*(2), 15−41.

Zhang, Z., & Zou, Y. (Eds.) (1953). *Dazhong tongjixue* [People's statistics]. Lixin kuaiji tushu yongpinshe chuban.

©2022 Arunabh Ghosh. This article is licensed under a Creative Commons Attribution (CC BY 4.0) International license, except where otherwise indicated with respect to particular material included in the article.

Footnotes

1. On the founding of the Econometric Society, see Bjerkholt (2017). For a list of past presidents, visit: https://www.econometricsociety.org/society/organization-and-governance/executive-committee-2021/past-presidents (retrieved February 17, 2022). ↩
2. At Nankai, Ho compiled 60 years of Chinese price and quantity data to calculate historical price indices, which he published in *Index Numbers of the Quantities and Prices of Imports and Exports and of the Barter Terms of Trade in China, 1867–1928* (Ho, 1930). For more on Ho and his relationship to Fisher, see Ho (1966, pp. 58–59). ↩
3. There is a large body of scholarship that traces the histories of probabilistic thinking, exemplary among which are Daston (1988), Gigerenzer et al. (1989), Kruger et al. (1987), and Porter (1995). ↩
4. This essay draws upon sections of my book, *Making It Count: Statistics and Statecraft in the People's Republic of China* (Ghosh, 2020); in particular, chapters 3 and 4. ↩
5. For a detailed discussion of these debates and their antecedents in the Soviet Union, see Ghosh (2018) and Ghosh (2020), in particular chapter 3. ↩
6. On the deployment of large-scale random sampling during the 1950s, see Ghosh (2020), in particular chapter 7. ↩

7. On these methods, and the challenges they generated, see Ghosh (2020), in particular chapter 5. ↩

8. Yang Hui, for instance, cited Winkler's (1954) essay "Older and Newer Ways of Solving the Index Numbers Problem" when mounting his critique of Fisher (Tang, 1956, p. 16). Winkler was among the earliest statisticians in the German-speaking parts of Europe to call for the application of mathematical and statistical tools to analyze—not merely describe—society (Schmetterer, 1984). ↩

9. The two other measures, the median and the mode, were deemed uncontroversial. ↩

10.
Named after their creators, Hermann Paasche (1851–1925) and Étienne Laspeyres (1834–1913), these are price and quantity indices:

$P_P = \frac{\sum P_n q_n}{\sum P_0 q_n}$ and $P_L = \frac{\sum P_n q_0}{\sum P_0 q_0}$. ↩

11. For more on these experiments, see Ghosh (2020, Part III). ↩